Synchrony within Triads using Virtual Reality

MARK ROMAN MILLER, Department of Computer Science, Stanford University, USA
NEERAJ SONALKAR, Center for Design Research, Stanford University, USA
ADE MABOGUNJE, Center for Design Research, Stanford University, USA
LARRY LEIFER, Center for Design Research, Stanford University, USA
JEREMY BAILENSON, Department of Communication, Stanford University, USA

Synchrony, the natural time-dependence of behavior in human interaction, is a pervasive feature of communication. However, most studies of synchrony have focused on dyadic interaction. In the current work, we explore synchrony in three-person teams using immersive virtual reality. Participants spent about two hours collaborating on four separate design tasks. The tracking data from the VR system allowed precise measurement of head and hand movements, facilitating calculation of synchrony. Results replicated previous work that found nonverbal synchrony in dyads in immersive VR. Moreover, we manipulated the context of the task environment, an informal garage or a traditional conference room. The environment for the task influenced synchrony, with higher levels occurring in the conference room than the garage. We also explored different methods of extending synchrony from dyads to triads, and explore the relationship of synchrony to turn taking and gaze. This paper provides theoretical insights about nonverbal synchrony and how design work functions in triads and provides suggestions for designers of VR to support good collaboration.

CCS Concepts: • **Human-centered computing ~ Collaborative and social computing ~ Empirical studies in collaborative and social computing** • Human-centered computing ~ Human computer interaction (HCI) ~ Interaction paradigms ~ Virtual reality

KEYWORDS: virtual reality, design teamwork, synchrony, work environment

ACM Reference format:
Mark Roman Miller, Neeraj Sonalkar Ade Mabogunje, Larry Leifer, Jeremy N. Bailenson. 2021. Synchrony within Triads using Virtual Reality. In *Proceedings of the ACM on Human-Computer Interaction,* Vol. 5, CSCW2, Article 400 (October 2021), 27 pages, https://doi.org/10.1145/3479544

1 INTRODUCTION

Collaboration is the means by which almost all human work is performed; yet for its importance and ubiquity, its processes are not fully understood [17, 50]. One process which occurs during collaboration is synchrony, the natural time-dependence of behavior in human interaction [14]. For example, two people engrossed in conversation may lean back in their seats at similar times. Alternatively, dyads can be "out of sync," and less likely to follow each other other's motions. In dyad studies, synchrony has been linked to task performance [1], prosociality [10], rapport

---

This.is work is supported by the Hasso Plattner Design Thinking Research Program for 2019-2020.

https://doi.org/10.1145/3479544

between therapists and clients [45, 46] and mother-child bonding [29]. The experimental manipulation of synchrony such as synchronized drumming [25] has led to increased cooperation [60].

The ubiquity of synchrony motivates the study of its causes and functions [28, 14]. Its broad definition has also led to a plethora of measurement techniques. The earliest works investigating synchrony [11] categorized the movement of body parts using video data and trained coders. Synchrony was operationalized in those studies as a covariation in behaviors occurring around the same time. More recent work [12, 46] has measured synchrony using Motion Energy Analysis [26], an automatic coding scheme based on pixel differences between frames of video for a coarse estimate of body motion. With more advanced techniques, such as computer vision [59, 42] or magnetic tracking [28], synchrony measures have included matching motion in individual body parts and facial expressions. Despite the heterogeneity of these measures, findings regarding synchrony have, in large part, been consistent across these methods of measurement.

In this work, we measure synchrony using virtual reality (VR). Virtual reality is a medium in which most of an organism's real-world sensory input is blocked out and replaced by computationally-generated sensory input. A key perceptual effect of this sensory input is the emergence of presence, the "perception of non-mediation" [35, 36]. When in a state of presence, the person's responses to the virtual stimuli are similar to actual stimuli. For example, when walking across a virtual plank over a virtual pit, a person tends to slow down and their hands begin to sweat [6, 56], responding as if the virtual danger of falling is real. Presence gives validity to the use of VR as an experimental method.

VR facilitates the measurement of synchrony by collecting motion tracking data. VR in use today depends upon accurate head and hand tracking for display and interaction, and this tracking data can be easily collected in a VR application. Once obtained, the data can be analyzed for behavioral signals [62] that, depending on the study, can range from left-right head rotations indicating anxiety [61] to interpersonal distance indicating implicit bias [16]. When the behavioral signal of interest is synchrony, the tracking data provides high temporal and spatial resolution of nonverbal synchrony, which is synchrony between nonverbal communication techniques like posture, head motion, and gesturing.

VR is not only a tool for accurately capturing participants' behavior, but it is also an immersive medium with the potential to computationally mediate interactions with self, others, and contextual environment [7]. In this work, we manipulate the contextual environment, whether participants complete their teamwork within a garage or a conference room. These locations were chosen as two places where design occurs, with the conference room being more formal and traditional than the garage. Based on the authors' experiences as design instructors, these contexts in the real world differ both in formality as well as the signaled tolerance for failure, leading to changes in how people approach their collaboration. We investigate the effect of these locations on team dynamics through synchrony.

This work contributes towards a deeper understanding of synchrony by focusing on triads, using VR to simulate avatars based in two different environments, and testing two different kinds of design tasks. Specifically, our contributions are (1) a study protocol for co-located VR studies using triads, (2) new methods for measuring triadic synchrony, and exploration of the properties of each method, (3) confirmatory results that synchrony occurs between teammates using VR, (4) confirmatory results that synchrony occurs among triads, (5) evidence that the context of virtual environment can affect synchrony, (6) evidence that speaking roles affect synchrony, and (7) evidence that gaze influences synchrony.

2 RELATED WORK

Our work is built upon related work in virtual reality and synchrony. Regarding virtual reality, we review work in multi-user VR and the effect of virtual environments on behavior. Regarding synchrony, we review studies of synchrony that use VR as well as synchrony in triadic interaction.

2.1 Multi-User Virtual Reality

Often owing to the technical difficulties in networking between virtual reality systems, empirical studies of communication and behavior in multi-user VR spaces [48] were limited for many years. This research gap is partially bridged based upon evidence that constructs such as immersion, presence, and behavioral realism extend effectively from individual and dyad studies into larger groups [30]. Work by Steed and colleagues [53] investigated the interaction of one VR user with two desktop users in a virtual environment and showed that the immersed user took a stronger leadership role in the collaborative puzzle-solving task. Schroeder and colleagues [51] also found an effect of the asymmetry of interfaces between immersive and desktop.

Recent technical developments in shared virtual environments have made multi-user VR more available to researchers and consumers. The emergence of social virtual reality platforms such as Altspace VR, VRChat, Mozilla Hubs, Oculus Rooms, Facebook Spaces, and High Fidelity has led to ethnographic work understanding the conceptualizations and usage of these platforms. One characteristic of current social VR is the interplay between grounding experiences in social context and adapting to the new affordances of the medium [38, 39]. For example, Dey and colleagues. [15] investigated the relationship between two VR users when one saw the field of view and heart rate of the other. The addition was a new affordance of VR, but it did not receive sufficient salience to affect the social experience. Latoschik and colleagues [34] produced an implementation of social VR using off-the-shelf hardware and software. They found that presence was higher in situations with 10 or more avatars than in situations with 5 or fewer, which suggests others in the environment can influence presence. Even with these insights, much work remains to be done in understanding and producing a good multi-user VR experience for all users.

2.2 Virtual Context and Behavior

Environments produced in VR have been known to affect user's behavior. Beginning with non-immersive virtual technologies such as Second Life, researchers have demonstrated that virtual environments can have effects like triggering stereotype threat, reducing test performance [8, 9]. Using immersive virtual reality, a virtual convenience store elicited smoking cravings from smokers even without any explicit cigarette stimuli [43], and college-age women showed higher body dissatisfaction in a virtual populated beach than in either an empty beach or party conditions [44].

Most studies of this type focus on an individual's behavior. However, there are few studies of multiple participants in immersive VR that investigate the link between context and some facet of interpersonal communication. A study by McCall and colleagues [37] involved both a manipulation of the virtual environment and networked participants, but the networked participants did not interact verbally. Rather, one participant was in the front of the room and the other was in the back while a virtual agent gave a persuasive speech, and the participants could see one another's head movements while they listened to the lecture. Results indicated students that were assigned to the front of the room were more persuaded and had more positive impressions of the speaker. In all, these studies show that a VR environment alone can cause significant changes in a user's behavior.

2.3 Synchrony in Virtual Reality

Evidence of synchrony carrying over to virtual reality has only recently been established. In work by Tarr and colleagues. [55] participants were asked to perform movements along with computer-controlled characters purported to be remotely located participants. The experimenters manipulated whether the agents covertly mimicked the participant's motion at a short (.25-.58s) or long (1.67-4.33s) delay. Participants who were covertly mimicked by the other characters felt greater social closeness to the characters than participants who were not mimicked did. However, this study did not investigate naturalistic synchrony, in which multiple participants interact in an everyday situation.

The first evidence of natural synchrony between multiple users in immersive virtual reality is work by Sun and colleagues [54]. In this study, 96 participants in 48 dyads completed an idea-generation task in virtual reality, framed as either a cooperative or competitive task. There was evidence of synchrony demonstrated by negative correlations between the amount of hand motion, interpreted as turn-taking in speaking and gesturing.

Gumilar and colleagues [27] investigated inter-brain synchrony in virtual reality in dyads. In inter-brain synchrony, the actions being compared across participants at the same time are the activation levels of regions of the brain. Two participants were instructed to perform a movement-matching exercise, intentionally mimicking the other participant's movement. Both conditions – real world and VR – led to significant brain activation synchrony between participants. Overall, these studies indicate VR is not a barrier to natural synchrony between people.

2.4 Relating Nonverbal Synchrony with Speaking Role and Gaze

In this work, we investigate the relationship of nonverbal synchrony with two other aspects of communication, speaker role and gaze. While both gaze and speaker role have been the subject of synchrony studies, they have functioned as a different operationalization of synchrony. For example, de Mendonça and colleagues used mutual gaze (two people looking at each other) and shared gaze (two people looking at the same object) as an operational definition of 'visual orientation' which they related to 'dyadic involvement' [29]. Walker and Trimboli [57] used the length of speaker transition times as indication of coordination between dyads, and operationalized synchrony as this speaking coordination. However, these works do not study nonverbal synchrony, i.e., the direction, location, and speed of body motion in conversation.

2.5 Synchrony in Triads

Investigation of synchrony in groups larger than dyads is sparse. There is some work dealing with synchrony among three people who are not peers, e.g., mother-father-child relationships [24] and couple-therapist relationships [13], but it is difficult to extend these findings to triads of peers. The most relevant works in this space are done by Wilkins and Nwogu [59] and Dale and colleagues [12].

In the study by Wilkins and Nwogu [59], video recordings of three-person conversations were processed using computer vision algorithms to extract body pose and facial action units. These values were combined in different ways to produce seven measures of synchrony, each of which correlated at some degree with human evaluations of synchrony. This data was then used to train a recurrent neural network to predict human-coded estimations of synchrony. This work demonstrated synchrony among three participants, but the focus of the paper was the automation of synchrony detection rather than exploring the various properties of synchrony among triads.

In the study by Dale and colleagues (12), motion energy analysis [26] was used to calculate synchrony between three participants. Results provided evidence for synchrony between each dyad within the triad, and furthermore reported exploratory results that the degree of synchrony between dyads within the triad was itself synchronized. While synchrony between dyads (rather than individuals) is not easily interpretable, especially considering one person is part of both dyads, it is an interesting finding and relevant to the larger question of synchrony in groups.

2.6 Current Work

The current work relies upon previous findings in VR, namely the effectiveness of VR as a medium for social interaction and as an influence on behavior through the virtual environment.

Overall, our focus in this work is a better understanding of synchrony. Specifically, we address three research gaps. First, we explore the effect of a virtual environment on team synchrony. To our knowledge, this is the first work that shows that location in which a team performs their work influences that team's nonverbal synchrony.

Second, we explore the extension of nonverbal synchrony to groups beyond dyads. In previous work, the interaction of each dyad within the triad is measured as its own unit and then aggregated into a team-level variable [59] or compared against other dyads [12]. We explore richer operationalizations of synchrony, including reinterpretations of existing measures, that may spark alternate conceptualizations of synchrony to be explored in future research.

Third, we follow the motivation of Dale and colleagues [12] and report relationships between synchrony and other variables we have recorded. In contrast to Dale and colleagues, though, we do not focus on team-level variables but rather moment-level variables, like turn taking and gaze. Ultimately, the investigation of these research gaps will lead to a better understanding of synchrony.

3 METHODS

3.1 Participants

Our study involved in total 21 teams. Of those teams, 5 had at least one participant not show up to the experiment. These teams were excluded from our analysis. Two more teams were excluded due to data collection errors; in particular, the tracking data timestamps were not sufficiently fine-grained to calculate synchrony. These restrictions reduced the number of valid teams to 14. Sessions were also thrown out when there was a long period of time (20 seconds or more) when one person was not producing tracking data, and therefore likely not visible. As a result, we were left with 44 sessions spread over 14 teams. Each team had at least one session in the conference room and one session in the garage.

Participants in this study had prior experience with design, evidenced either by taking design courses or working in a professional designer capacity. The degree of each participant's VR usage was not collected at the time of the experiment. However, during debriefing, some participants did share they had prior experience in virtual reality, though none had experience with the specific VR platform we chose.

3.2 Apparatus

Because the content of this study is both virtual and physical, the apparatus of this study has multiple sub-sections. The apparatus consists of the physical space, the hardware interfacing with the physical world, the software running during the experiment, the virtual content displayed to the participants while in VR, and the design prompts used as study conditions.

3.2.1 Space

The physical space for these experiments consisted of two adjacent rooms in a campus building associated with the design department. The larger room was 5m by 3m and the smaller was 4m by 3m. In order to have walking space (about 1.5m by 2m) for each of the three participants, the larger room was separated into two by a blackout curtain. The necessity of the blackout curtain came from previous experience having multiple HTC Vive VR systems in the same area. Because the Vive Lighthouses use infrared light to track position, it is possible for two tracking systems to interfere with each other. The curtain was kept shut while participants were in VR, blocking this light. Other tracking errors also occurred due to a wide-open door between the two rooms, which was addressed by keeping the door shut, and a highly reflective large glass pane on one of the walls, which was addressed by covering it in paper.

Within each participant's space, there was also a physical table and chair. Though participants were standing during the work sessions in VR, participants sat at the table between sessions to fill out the study questionnaires. While participants were in VR, there was one researcher in each room whose job was to 'spot' the participant, ensuring the participant did not run into a wall or hit a table. A schematic of the space is given in Figure 1 panel A, and a photograph of one room is in Figure 1 panel B.

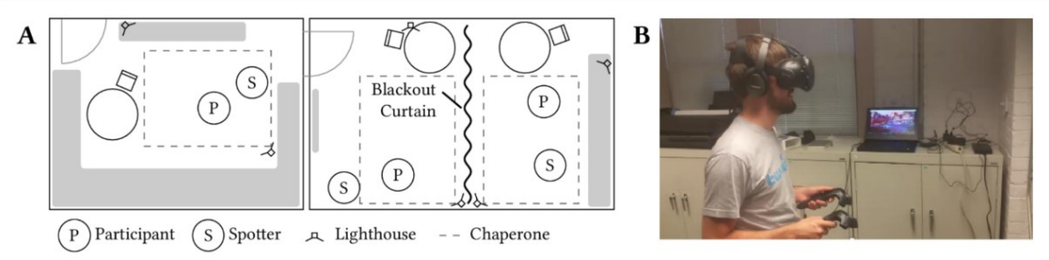

Figure 1. Experiment Space. In panel (A) is a schematic of the experiment space, consisting of two rooms, one was split with a blackout curtain in order to fit two participants. Icons are given for participants, spotters, Vive Lighthouse devices, and the Vive Chaperone border. Light gray areas indicate space taken up by other objects (for example, desks, monitors, and cabinets).

The spaces were chosen to be adjacent to help the logistics of running the study. Co-location ensured researchers could communicate quickly and easily, and could share solutions to technical problems. One unexpected consequence of adjacency was that sometimes participants could hear the other participants in through the physical space rather than only in VR. This problem was solved with the addition of over-ear noise-cancelling headphones for each participant.

3.2.2 Hardware

The virtual reality headset in use was the HTC Vive headset and controllers. This headset was chosen for its ability to easily facilitate and track movements in a room-scale space. These headsets were connected to one VR-enabled laptop computer for each of the three participants. A fourth computer acted as a server hosting the virtual environment and virtual content. Originally, we had one of the VR computers host the virtual environment, but that computer kept crashing during the experiment. Headphones were used both to render the virtual sound but also to block out noises from the physical world. Three camcorders, one in each VR space, recorded both audio and video of the participant in the physical world.

3.2.3 Software

The software used for rendering and networking the multi-user VR space was High Fidelity, a consumer facing product that was available at the time of the experiment. This platform was chosen because it was flexible enough to allow collection of tracking data while still providing simple and effective multi-user VR networking. The company has since ceased development on this project [49] and opened the code base to the public [63], and the project has been continued as an open-source project by others [33].

The tracking data was transmitted from the High Fidelity application to a Django 2.0 server running in Python 3.6 hosted on Heroku, which stored the tracking data within a PostgreSQL database. The first-person VR audio and video were captured using OBS Studio [41], an open-source software for screen recording.

3.2.4 Virtual Content

There were three virtual locations the participants experienced: a garage, a conference room, and a 'waiting room'. All participants began their VR experience in the 'waiting room', which was not designed to resemble a space encountered in the physical world but rather encourage familiarity with VR as a medium. The waiting room was a large room (15m x 15m) containing a mirror and a few geometric shapes. The mirror was placed to allow participants to see their avatars and begin to associate the avatars with their own bodies [20, 52]. The geometric shapes were physics-enabled, so participants could stack them and play catch. We included these objects to encourage low-threshold interaction between participants [38].

In the garage, there was a whiteboard on two-by-fours, an old couch, and some car tools in the corner. A car was in the driveway outside, and the background showed a residential neighborhood. In the conference room, there was a whiteboard mounted to the wall, and similar-looking conference rooms were across the hall. A floor-to-wall window showed a city panorama. Figure 2 shows screenshots of these two locations.

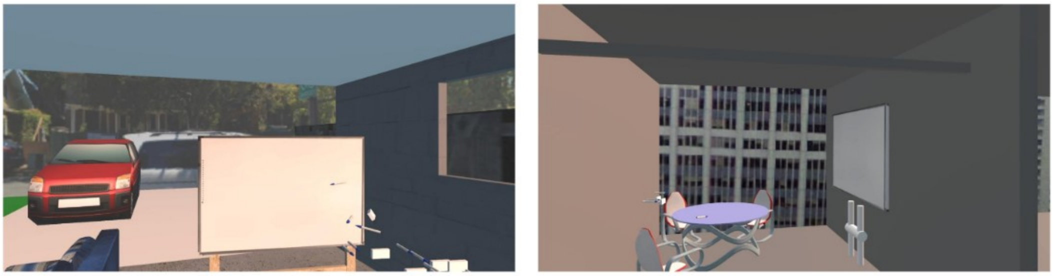

Figure 2. Screen captures of the two locations of the study: the garage and the conference room.

Characters were selected by participants, and were limited to avatars given in High Fidelity. For more flexibility in avatar representation, there are avatar customization tools such as Adobe Fuse; however, we had difficulties integrating the mouth motion of these models into the software. We judged the visual cue of mouth movement and its role in signaling who was speaking would be more important to interpersonal communication than the avatar. Participants could see and hear each other in the virtual environment. Locomotion in VR could be achieved by using real body motion, controller button presses, or a teleporting mechanic. Participants could also draw in 3D space using a virtual marker and eraser. Screenshots of the virtual world in action are given in Figure 3.

Figure 3. Screen captures of each participant's view during the session.

3.2.5 Design Tasks

The task prompts were written by the second author and had been piloted in face-to-face meetings as part of course material. The first decision-making task was to choose between three new wearable product prototypes to pitch to the CEO and CTO of the readers' company. The second decision-making task was deciding the target population for beginning development of incubators. The first concept-generation task was brainstorming new personal mobility options for college students on campus. The second concept-generation task was brainstorming high-end accessories to complement self-driving cars. The full text of the design prompts are available in the supplemental material. Participants read the written design prompt on a physical sheet of paper, and then entered one of two VR locations: garage and conference room.

3.3 Procedure

Upon arriving, participants signed a consent form approved by the university IRB. When all three participants had arrived, the researchers led the participants to their respective virtual reality rooms. Participants began in the mirror room to accustom themselves with the virtual reality controls. This first phase lasted about five minutes. This time also gave researchers leeway in addressing any technical or setup issues.

Participants then had four sessions working on design tasks as a team. In each session, the participants began outside of VR in a physical chair in the experiment room and read a prompt of either a concept generation or decision making task. The prompts are available in the supplementary material. Once all participants were familiar with the prompt, they re-entered VR at the virtual location proper to the session: either the garage or the conference room. The participants were given ten minutes to complete each task. At the end of the task, the participants stepped out of VR and completed the task questionnaire. After all four sessions were completed, participants gathered in a single room in person to reflect upon and discuss the VR experience.

3.4 Variables

3.4.1 Independent Variables

The two independent variables in this study are the virtual setting and task type. The two virtual settings are the conference room and the garage, shown in Figure 2. These were chosen based on the authors' experience and observation as design instructors for their cultural connotations of formality and tolerance for failure. We chose two kinds of tasks for this study -

concept generation tasks and decision making (also known as concept selection) tasks. These tasks instantiate the divergent and convergent phases of design work [23].

The design tasks were generated with the following four requirements in mind: the task must be realistic enough to engender motivation in the participants to engage and complete the task; the task must be simple enough to be completed in 10 minutes; the task must fit the virtual environment - either conference room or garage - setting in which it was to be conducted; and the task must call on the design training of the participants, but must be general enough that a design thinker with no specialized engineering background could complete it.

3.4.2 Dependent Variables

The dependent variable of interest in this study is nonverbal synchrony. This is calculated based upon the head and hands position data collected by High Fidelity, and measures the degree to which the participants' amounts of motion correlate over time. Nonverbal synchrony can be computed upon a dyad's motion data, or upon a group's motion data collectively. A more precise formulation of this value is given within sections 3.5.2 to 3.5.6 and the supplementary material.

In addition to the data-intensive computational methods, there were questionnaires given between each session. These questionnaires included questions on Inclusion of Other in Self [2], visual reflection activities [31], and two questions asking the participant's judgement of the effectiveness of the team relative to either previous sessions with the same team or all previous team experiences.

Team performance was measured differently between the concept generation and decision making tasks. The concept generation tasks were judged on the quantity and novelty of ideas. However, because the decision making task cannot be judged upon a 'distance' to some 'true' answer, we could only measure it based on a decision process metric. We decided to use the Amplex-Limit Process framework [18]. In this framework, there is one type of utterance, 'limit handling', which has been correlated to the quality of a decision in previous work. A 'limit handling' operation is when a team member "implicitly or explicitly questions, challenges, qualifies, or conditions a limit." In turn, a limit "excludes a concept or set of concepts from discussion based on attributes of the concepts." In previous work [18], teams with more limit handling operations were rated as higher-performing, so we use this to base our measure of team performance.

3.5 Analysis

The analysis to produce values indicating synchrony consists of several steps. In short, we describe the tracking data, choose a measure of synchrony, and perform an inferential test. However, there is a fair degree of complexity underneath this process.

3.5.1 Tracking Data.

The collected tracking data consisted of many samples, each of which contained 3D position data (X, Y, Z) and 3D rotation data (formatted as a quaternion) for both the participant's head (as tracked by the headset) and hands (as tracked by the hand controllers). This data was programmed to be collected at the rate of 20Hz. Data analysis was performed in R version 4.0.0 using packages from "tidyverse" [58] and tools for 3D data from "dddr" [40].

For most of the data, i.e, 99.6% of samples, they were collected reliably (within 18-22ms of the previous one). However, the duration between samples could be exceptionally large, if, say, one participant had a technical malfunction and disconnected from the environment. In that case, if there was more than 20 seconds without tracking data from a participant in the triad, we rejected the tracking data.

3.5.2 Measuring Synchrony

Our primary measure for this analysis was motion synchrony. There are many ways motion synchrony has been calculated, e.g., [3, 21, 22, 54]. Synchrony based on velocity data [12, 22, 42, 46, 54, 59] appears to be more common in the literature than synchrony based on rotation data [28, 61]. Of the three points tracked by a VR headset, the head is the one that most closely tracks gross rigid body motion, the primary component of body motion [22]. Rank correlation is chosen over regular Pearson correlation because head speed is not normally distributed over time. Therefore, the measure of dyadic synchrony we use is Spearman correlation of head speed. While computing synchrony this way involves fewer sources of streams than some previous work, the current heterogeneity in synchrony measurement indicates synchrony is a construct robust to many kinds of measurements. Moreover, understanding synchrony upon merely head and hands is important as VR systems become more popular.

3.5.3 Measuring Synchrony in Triads

One of the innovations of this work is the extension of synchrony into groups larger than two. To speak more precisely about these measures, we will define some mathematical expressions. Let m_{pt} represent the motion (in this work, head speed) of participant $p \in P$ at time index $\mathrm{t}, 1 \leq \mathrm{t} < \mathrm{s}$. In this notation, the dyadic synchrony between a and b is $\mathrm{cor_t}(\mathrm{m_{at}}, \mathrm{m_{bt}})$, with cor_t indicating correlation over time.

One simple approach to extend synchrony to larger groups is to calculate the synchrony scores among each dyad (three dyads in our three-participant case) and then average them. This approach appears in work by Wilkins and Nwogu [59] under the names "AllAUCorr" and "3AUCorr". We name this measure $Sync_{avg}$ and calculate it formally as:

$$\mathrm{Sync_{avg}}(P) = \frac{1}{|P|(|P|-1)} \sum_{p \in P} \sum_{q \in P, p \neq Q} cor_t\left(m_{pt}, m_{qt}\right)$$

However, this method seems to assert groups are merely collections of dyadic interactions, which is a questionable assumption at face value. To address this, we also consider a second approach. Intuitively, if the motions of different participants "track together" then there will be less variance in motion at a given time than across all times. After ranking and normalizing the motion values $\widehat{m_{pt}} = \frac{m_{pt} - mean_t(m_{pt})}{sd_t(m_{pt})}$ relative to each participant, the variance across all participant samples at each time index is computed, producing a variance for each time index t. Then, all of these variances are averaged over time. This averaged variance is then normalized by dividing by the variance of all samples over the full length of time. Note that high synchrony would be associated with low variance, so we define the output measure $Sync_{tvar}$ as one minus the quotient mentioned above. This expression is written as:

$$\mathrm{Sync}_{tvar}(P) = 1 - \frac{\frac{1}{s}\sum_{t=1}^{s} var\left(\{\widehat{m_{pt}} \mid p \in P\}\right)}{var\left(\{\widehat{m_{pt}} \mid p \in P, 1 \leq t < s\}\right)}$$

While these two measures of synchrony seem to have different interpretations, we show in this work that they are in fact equivalent up to a scaling factor and a constant, namely:

$$\mathrm{Sync}_{tvar}(P) = \frac{|P|-1}{|P|} Sync_{avg}(P) + \frac{1}{|P|}$$

This fact can be derived from the law of total variance, but a longer yet more elementary proof of this for teams of any size is given in the supplemental material. This convergence of approaches, which is a novel result in regard to the study of synchrony, provides some degree of confidence in the usefulness of this measure in either of its forms.

Even though the values are linearly equivalent, we still must choose one value to work with. In this decision, it is worth noting that a $Sync_{avg}$ of zero has a meaningful interpretation that there was 'no synchrony,' i.e., $Sync_{avg}$ among multiple time series that are independent of each other is zero. In contrast, $Sync_{tvar}$ among independent time series is $\frac{1}{|P|}$. This fact, alongside the use of $Sync_{avg}$ in previous work [59], leads to our conclusion to use $Sync_{avg}$ as the synchrony measure in this work.

3.5.4 Measuring Synchrony over Time

It is also of interest to consider what periods of time contribute to synchrony. This can enable reasoning about the conditions and even causes of synchrony. A naïve approach is to calculate the correlation for a specific time interval. However, synchrony measured in this way cannot be composed and decomposed, e.g., the correlation between two participant's motions during a minute cannot be calculated merely given the correlation within each second of that minute.

The compositional measure of synchrony that we propose to use here is the *z-score product.* In this operation, the variable of interest (e.g., motion) is z-scored relative to the entire duration, the two variables are multiplied elementwise, and finally the average is taken across the time period of interest. Using the notation from sections 3.5.2 and 3.5.3, the expression we propose for synchrony within a duration of interest $\mathrm{D} \subseteq [1, s]$ is

$$\mathrm{Sync}_{\mathrm{dur}}(P, D) = \frac{1}{|P|(|P|-1)(|d|)} \sum_{p \in P} \sum_{q \in P, p \neq Q} \sum_{\mathrm{t} \in \mathrm{D}} \widehat{\mathrm{m}_{\mathrm{pt}}} \widehat{\mathrm{m}_{\mathrm{qt}}}$$

When the time period of interest is the entire source, i.e., $d = [1, s]$ we recover the original correlation-based measure of synchrony, as the correlation coefficient of two variables can be written as the average elementwise product of the z-scores. Multiple time durations D_a and D_b can be composed by taking the average of their z-scored product synchrony scores, weighted by their duration using the formula

$$Sync_{dur}(P, D) = \frac{|D_a| \cdot Sync_{dur}(P, D_a) + |D_b| \cdot Sync_{dur}(P, D_b)}{|D_a| + |D_b|}$$

This compositional measure is useful for two of our analyses: synchrony during speaker turns and synchrony and gaze.

3.5.5 Measuring Synchrony Relative to Speaker Turns

One potential variable that may influence synchrony is the role of the speaker. In this work, we investigate two possible relationships the speaker role may play regarding synchrony. First, we investigate whether the coordination required to manage turn-taking effectively manifests itself in synchrony. For example, consider two people, Alice and Bob, speaking together. When Alice finishes her turn, Bob replies to her. There is coordination between Alice and Bob that is necessary for Bob to be ready to generate his utterance quickly. It is possible this coordination appears as synchrony between Alice and Bob around the time Alice's turn ends and Bob's begins.

The second relationship we consider is whether during an utterance the speaker-listener dyads have higher synchrony than listener-listener dyads. This prediction is reasonable if synchrony depends on attention, and both listeners are paying attention to the speaker rather than to each other. A graphical description of these two methods is provided in Figure 4.

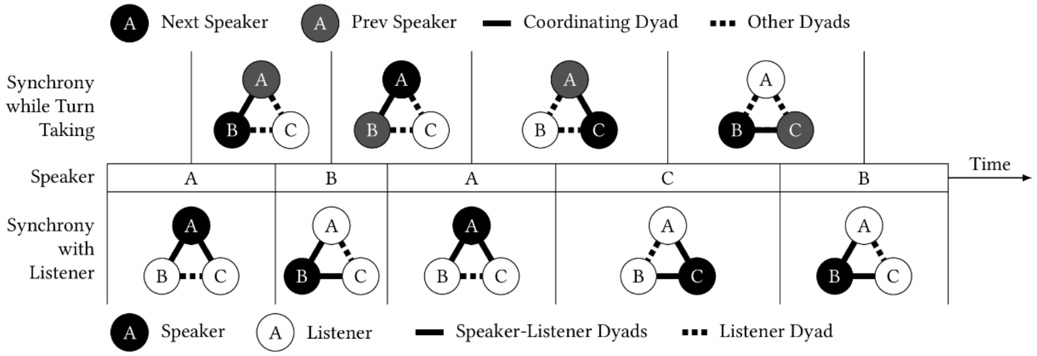

Figure 4. Schematic of the two measurements of synchrony based on speaker turns. Time flows left to right; above and below the center line are the two ways we sort moments of dyadic interaction into types.

To determine the speaking turns, verbal interaction information was produced from first-person video and audio data. For each participant, the recording software OBS Studio recorded the VR view of the participant along with the microphone input (from the self) and VR audio output (from teammates). To produce a coherent recording of the interaction, these video streams were manually aligned, and one recording's audio was chosen to accompany the three streams. Not all three audio streams could be used, because they had slight differences in timing, which made listening to the combined audio streams unpleasant. Then, a trained coder annotated the combined recording. The output of this annotation process was a set of speaking turns (and silences) with the corresponding start time and speaker of the turn.

3.5.6 Measuring Synchrony Relative to Gaze

Continuing with the thread of synchrony as attention, one may expect synchrony to be higher when participants are looking at each other compared to when they are not. Because we have collected head position and rotation data, it is possible to estimate how close a target participant is to the center of the looking participant's field of view. The measure we use is the angle between the forward vector of the looker's headset and the vector from the looker's headset to the target's headset. Note that this is a measure of head direction, not necessarily gaze direction. When this value is zero degrees, the target is in the center of the looker's field of view. Note that this measure is adirectional: thirty degrees above, below, to the left, or to the right of the center of the field of view all produce a value of thirty degrees.

There are two more decisions to consider in operationalizing gaze. First, what is the threshold between "looking" and "not looking?" While head direction does not necessarily give gaze direction, there is usually less than 15 degrees between head direction and gaze [19]. One possible value is a threshold of 15 degrees. Another potential value is drawing a boundary at 70 degrees, which defines as "in view" any position within the field of view of the headset. Because this work is exploratory, we report a range of thresholds varying from 0 to 90 degrees.

Second, what happens if one participant is looking at the other, but not vice versa? One may claim that "looking at" is one-directional: it is important only to know if A is looking at B, OR B is looking at A. Another may claim that "looking at" needs to be two directional: it is important two know if A is looking at B AND B is looking at A. Again, because this work is exploratory, we report both values, using the terminology AND and OR for these two ways of combining the values. Combined with the range of thresholds, this work begins to probe the relationship of nonverbal synchrony and gaze.

3.5.7 Pseudosynchrony

Because measurement of synchrony is complicated, we discuss the significance tests commonly used in its analysis. Because time samples from the same participant are not independent, one should not report a significant correlation between the motion of two participants as evidence of synchrony. Instead, statistical significance is measured using the paradigm of pseudosynchrony [5], where true correlation is compared to correlation between time-lagged dyads. Pseudosynchrony is estimated by drawing samples of synchrony from the null distribution, i.e, under the assumption that time alignment is not a factor. A portion of motion of person A is correlated to the motion of person B some time later. If synchrony is occurring, then regardless of the distribution and dependence of motion over time, there should be a stronger relationship between real motions than pseudosynchronous motions. In all tests of the existence of synchrony, we report results compared against pseudosynchrony.

3.6 Hypotheses and Research Questions

To organize the results of our work we will refer to hypotheses through the results section. The hypotheses are differentiated into hypotheses H1 and H2, which were tests considered from the outset validated by some previous work, and RQ1-5, which are more exploratory results with less grounding work.

- H1: Synchrony occurs among triads in virtual reality.
- H2: Synchrony will correlate with task performance.
- RQ1: Will synchrony differ across sessions based upon environment?
- RQ2: Will synchrony differ across sessions based upon task type?
- RQ3: Will dyads coordinating the speaker role be more or less synchronous than the other dyads?
- RQ4: Will listener-speaker dyads be more or less synchronous than listener-listener dyads?
- RQ5: Will synchrony be higher or lower when pairs are looking at each other than when they are not?

4 RESULTS

Guided by our hypotheses and research questions, we report the results of several analysis. We begin with the most general, evidence for synchrony, and then explore synchrony by condition and task, by task performance, by triad, by speaking behavior, by gaze behavior, and finally discuss novel measures of group synchrony.

4.1 Synchrony

There was strong evidence of synchrony among participants. By comparing synchrony of participants' motion with the null distribution of pseudosynchronous interactions, a significance level can be estimated. Three thousand pseudosynchronous interactions were generated for each session. The mean synchrony (calculated as Sync_{avg} above) across all real sessions (M = 0.0455) was compared to the distribution of the mean synchrony for 44 pseudosynchronous sessions (M = $-3.76 \times 10\text{-}5$, SD = 0.00414). No mean of the 3000 simulations produced by pseudosynchrony was larger than the observed mean, so we can conclude strong evidence for synchrony among participants in a team, $p < 0.001$, providing evidence for H1.

In figure 5 is a descriptive plot of synchrony, visible as synchrony dropping as each dyad's time series are shifted out of alignment, usually only taking a few seconds of shift to reduce the effect sizably.

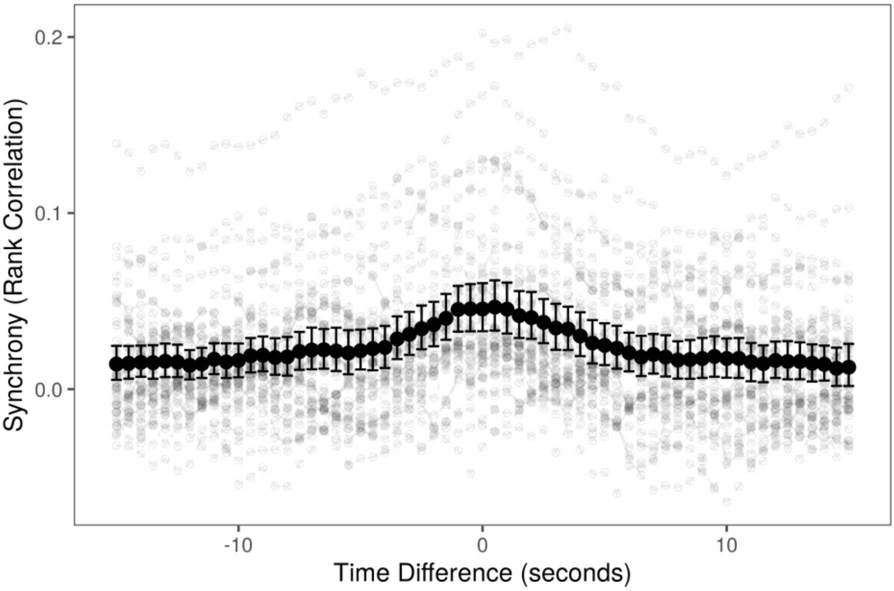

Figure 5. Synchrony and Time Offset. Synchrony is highest when interactions align; if the datastreams are shifted slightly in time (by 1-5 seconds) the correlation becomes much weaker. Translucent points represent team synchrony; solid black dots represent means; error bars indicate 95% bootstrapped confidence intervals of the mean.

4.2 Synchrony and Task Performance

Synchrony did not show a statistically significant relationship with performance in either the concept generation tasks or decision making tasks. In a mixed-effect model predicting total number of ideas generated from synchrony, with a random effect for each team, the effect was not significant, t(12.4) = 0.485, p = 0.636. In a mixed-effect model predicting number of limit handling events from synchrony, with a random effect for each team, the effect was not significant, t(14.5) = 0.936, p = 0.365. These results do not support H2.

4.3 Synchrony Between Conditions

We also found evidence of an effect of virtual location on synchrony. To show this, we first ran a mixed-effects model predicting each dyad's synchrony score based on location, and including a random effect for each team and for each session nested within the team. The result was a singular fit with the random effect of team equal to zero. This can be interpreted as team accounting for very little of the variance in synchrony, once the session variance is accounted for. We dropped the random effect of team but preserved the random effect of session. The difference between conference and garage locations was significant, $t(36) = -2.32$, $p = 0.0261$, answering RQ1 such that there were higher synchrony scores among sessions within the conference room than sessions within the garage.

There was no evidence for a difference in synchrony due to task (decision making and concept generation). As before, we use a mixed-effects model, and we did not find a significant difference, $t(36) = 0.092$, $p = 0.92$. providing no evidence for either side of RQ2.

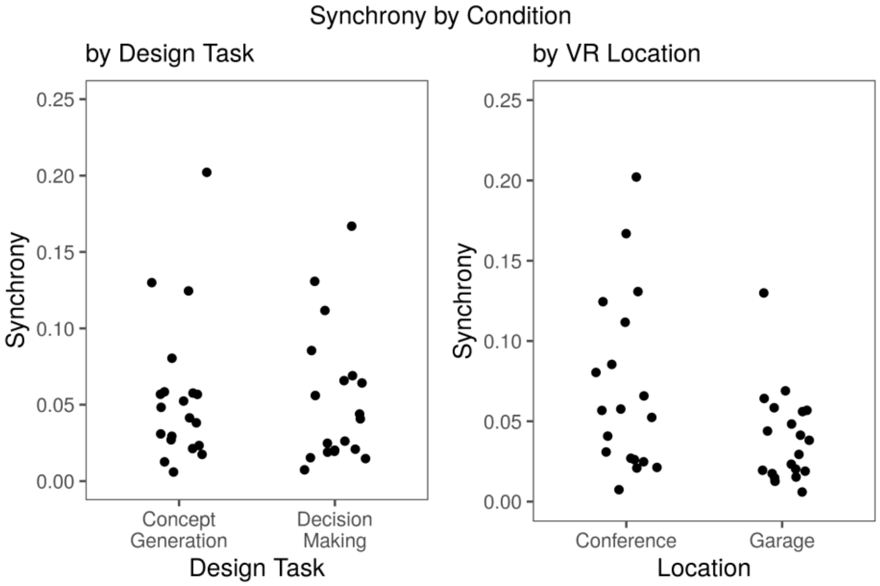

Figure 6. Synchrony scores per trial are displayed on points for each of the two comparisons. Synchrony is significantly higher in the Conference condition than the Garage.

4.4 Synchrony in Triads

As this is one of few studies investigating synchrony among three participants, we report descriptive analyses of working with triadic synchrony. The plots in Figure 7 are a plot of all synchrony scores over all teams and tasks. While there is variations in synchrony among dyads, there is also a mild effect of session on synchrony, i.e., synchrony among dyads in the same triad tend to be related. In a one-way ANOVA, session (Task + Team) predicted synchrony, F(43) = 2.65, $p < 0.001$. We interpret this value with some caution, as synchrony between two dyads in the same triad are not independent, both being dependent on the motion of the person common to both dyads. However, it is also not clear how to account for this fact, and we report this result with this caveat.

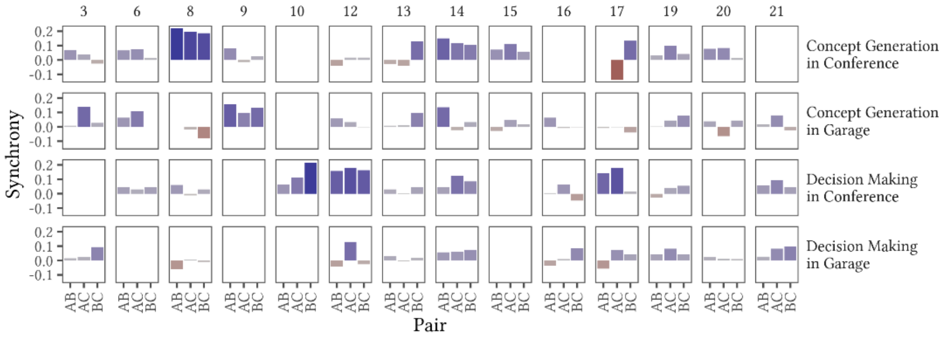

Figure 7. Synchrony of each dyad in the triad broken down by team and task. Blank rectangles indicate sessions with missing data (section 3.5.1).

The difference in synchrony across locations (rows 1 and 3 vs. rows 2 and 4) is somewhat visible in this plot. However, this plot is more interesting in showing relationships of synchrony among dyads in a triad. It is possible for all three triads to have high synchrony (Team 8, concept generation in conference room; Team 12, decision making in conference room), for one dyad to have much lower synchrony (Team 17, concept generation in garage), and all nearly no synchrony (Team 20, decision making in garage). Prompted by the extension of synchrony to triads, future work can explore synchrony of dyads or subgroups within larger groups.

4.5 Synchrony and Speaking

One way to probe the relationship between verbal and nonverbal behavior is to focus in on the relationship between synchrony and turn-taking. Does the pattern of turn taking, e.g., who is the speaker, who is transitioning to be the next speaker, relate with our measures of synchrony?

We explore this question by evaluating some methods to predict which dyad among the three in a triad has synchrony that is reliably larger or smaller than the others. This prediction can be done at many levels, but for our initial explorations in this work, we choose two: predictions can be made based upon turns or sessions.

4.5.1 Synchrony in Dyads at the Turn Level.

There are two analyses performed at the turn level. The difference is whether the unit of analysis is turns or transitions. Transitions are defined as the periods of time in between the middle points of sequential turns, as described in section 3.5.5 and figure 4.

In each of the two analyses, the primary model is a mixed-effects model predicting synchrony, as measured by z-score product, from the type of dyad (speaker-listener / listener-listener, coordinating / other) with a random effect of team. The model weighted each data point based upon the length of the time segment such that turns or transitions that were longer (and therefore had less variance) were weighted more strongly.

The first comparison is between coordinating and non-coordinating dyads. Recall that coordinating dyads manage a transition, i.e., dyads where one person stops speaking and the other starts. For each transition, there are two dyads that are not transitioning and one dyad that is. Whether the dyad was transitioning or not did not predict z-score product synchrony ($t(5232) = -1.192$, $p = 0.233$), providing no information related to RQ3.

The second comparison is whether the synchrony scores of the two dyads of the triad that involve the speaker in each turn are reliably different from the third dyad that does not involve the speaker. We found a significant relationship between whether the dyad involved the speaker and dyad synchrony ($t(5814) = -4.597$, $p < 0.001$) such that dyads between non-speakers had higher synchrony (0.071) than dyads involving the speaker (0.031). This result answers RQ4 that speaker-listener dyads do have less synchrony than listener-listener dyads.

4.5.2 Synchrony in Dyads at the Session Level.

In addition to exploring synchrony by turns, we explored whether some feature of a dyad's verbal behavior over the whole session predicted synchrony, namely the degree to which the dyad talked together, as measured by the number of transitions where both ends were filled by the two participants.

In a mixed effect model predicting a dyad's synchrony based upon total shared transitions, and a random effect of session nested within team, total shared transitions did not show significance ($t(75.06) = 0.400$, $p = 0.690$).

4.6 Synchrony and Gaze

Another important feature of conversation is gaze. In this work, we ask how synchrony changes depending on whether two participants are looking at each other. This result depends on what kind of behavior is considered “looking.” We first report some descriptive values of gaze within triads in order to better ground the findings that follow relating gaze to synchrony. These are described in Figure 8. Then, we report the z-score product measure of synchrony conditioned on gaze behavior. Finally, we explore a potential confounding factor in a follow-up test.

4.6.1 Gaze Angle Distribution

To ground our study of synchrony and gaze, we first report descriptive statistics of gaze, namely, the cumulative distribution of gaze angle, minimum gaze angle within a dyad (OR), and maximum gaze angle within a dyad (AND).

In Figure 8, there are three panels. Each panel plots the cumulative distribution function of gaze angle over different conditions, showing the percentage of samples considered “looking”, that is, less than the threshold specified on the x-axis. When the threshold is low, most samples are counted as “not looking” and so the percentage of time spent looking is low, and when the threshold is high, the percentage of time spent looking is high.

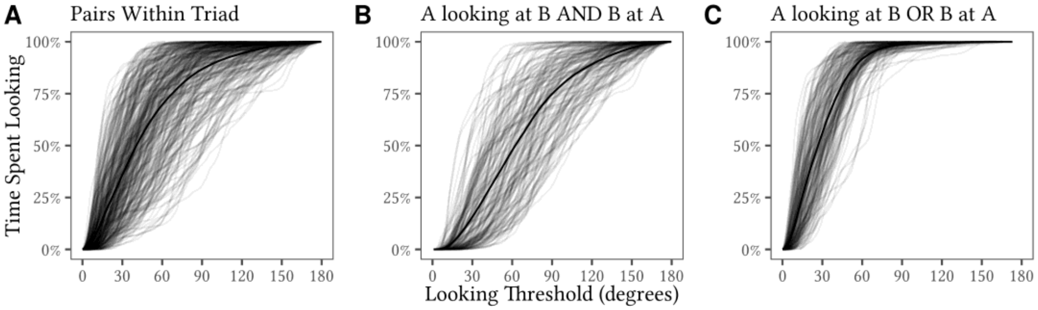

Figure 8. Cumulative distribution functions of angle between the forward vector of the looker’s head and the vector from the looker’s head to the target’s head. Panel A shows two lines per dyad because gaze angle is bidirectional. Panels B and C both focus on dyads but use either the lesser (OR) or greater (AND) of the two angles at each sample. Black lines indicate the marginal cumulative distribution across all dyads.

In panel A, one line is displayed per looker-target pair. There was wide variation among dyads in how often a looker looked at a target. Considering the vertical slice of data where the looking threshold is 60, one can see there was a looker who had the target within 60 degrees of center only 12 percent of the time, but there was another looker who had the target within 60 degrees of center nearly the entire time.

In panels B and C are plots of time spent looking where both directions are considered, that is, whether team member A was looking at B and/or vice versa. For example, in panel B, an x-value of 15 degrees indicates time spent looking is, on average, 5%. Interpreting this, one can conclude that for randomly chosen teammates A and B, A is within 15 degrees of center of B’s field of view and B is within 15 degrees of center of A’s field of view only 5% of the time.

In panel C, an x-value of 70 degrees, which encompasses the FOV of the VR headset, indicates time spent looking is almost 90%. This means that for randomly chosen teammates A and B, 90% of the time, either A can see B in the headset or B can see A in the headset. In sum, having concrete values when thinking about gaze angle provides a better context for understanding the relationship between synchrony and gaze.

4.6.2 Synchrony Conditioned on Gaze Behavior

Ultimately, the interest we have in this work is how gaze might be related to synchrony. A reasonable hypothesis is that participants who can see each other will have higher synchrony. We continue with the options described in section 3.5.6, namely the angle threshold (0-180) and the function combining the two conditions (AND, OR).

In figure 9, there are two panels, comparing synchrony when participants are and are not looking at each other across all these parameters. The left panel uses the AND operation, i.e., the red curve consists of synchrony scores between participants A and B when A is looking at B and B is looking at A. Across a range of threshold values, moments when both participants are looking at each other (red) are lower in synchrony than moments when at least one participant is not looking at the other. The right panel uses the OR operation, i.e., the red curve consists of the synchrony scores between participants A and B in moments when A is looking at B or B is looking at A. Again, synchrony appears to be lower when one participant is looking at the other. These results answer RQ5 that gaze behavior does relate to synchrony.

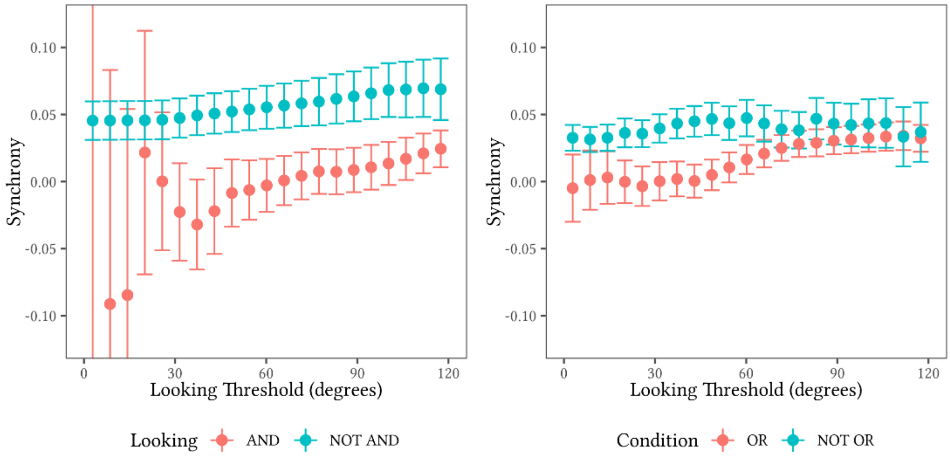

Figure 9. Synchrony and gaze. Across a range of parameters, participants tend to have higher synchrony when not looking at each other. Each dot represents the mean synchrony across each dyad and session, which in turn was calculated by the z-score product defined in section 3.5.4. Error bars are 1.96 times standard error of these scores based on dyads as a unit of analysis. The error bars on the left side of the AND condition are large because there are few samples in which both participants look at each other so directly (see figure 8).

Why might this be the case? While we do not have a clear, confident answer for this result, we do wish to report one analysis to inform future work. In short, we speculate the reason is synchrony is based upon motion, and motion influences whether someone is looking at someone else. Figure 10 contains a plot showing the chance a looker is looking at a target conditioned upon the percentiles of the looker's and target's head speeds.

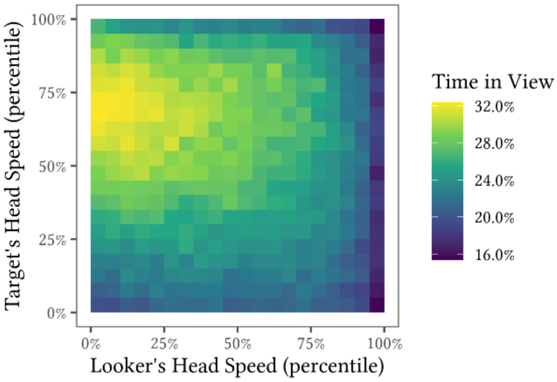

Figure 10. Given the head speed of a target and a looker, what is the likelihood the looker is looking at the target?

The type of motion in which the looker is most likely to be looking at the target, with a likelihood of about 30%, is when the looker's head speed was very slow (<15th percentile, about 0.5-1cm/s) and the target's head speed was slightly faster than the median (65-75th percentile, about 5-7cm/s). This speed can be interpreted as one participant taking animatedly or walking slowly. In contrast, the moments in which the looker is least likely to look at the target, with a likelihood around 15%, are when the looker's head speed is very high (>90th percentile, >12cm/s). This condition can be interpreted by noting that when a participant is moving they are likely looking where they are going rather than at another person.

The moments with the highest synchrony are plotted in this graph in the top right and bottom left quadrants, I.e., when both participants' head speeds are high or low together. The moments where a participant is more likely to look at another is in the top left quadrant, i.e., when one is moving and the other is not. This may have led to the difference in synchrony depending on visibility. While it would be ideal to account for this dependency between speed and looking in some way, we have not found it straightforward and so we have left it for future work.

In summary, we have found that synchrony tends to be lower when one participant is looking at the other. While the mechanism is unclear, it is likely influenced by the fact that motion and looking are related.

4.7 Alternative Measures of Triadic Synchrony

In the extension of synchrony from dyads to triads, operationalization of synchrony becomes more complex. We consider this added complexity to be a benefit, as it provides a testing ground to compare against different conceptualizations of synchrony. In this section, we introduce four additional measures of synchrony and compare them using the same data we have collected.

One potential measure is motivated by the fact that human attention is limited, and participants may only be able to be 'in sync' with one participant at a time. Therefore, synchrony at a moment in time may be better represented by the *maximum* synchrony over the three dyads rather than the average.

Producing a mathematical expression for this intuition presents a new potential variable. At what point is the maximum applied? For simplicity's sake, we restrict ourselves to two options. First, the maximum can be applied once synchrony is calculated over the entire duration of the session, but within each dyad separately, leading to the expression for "session maximum"

$$\mathrm{Sync}_{sesmax}(P) = \max_{p,q\in P, p\neq q} \mathrm{cor}_t\left(m_{pt}, m_{qt}\right) = \max_{p,q\in P, p\neq q} \frac{1}{s}\sum_{t=1}^{s} m_{pt} m_{qt}$$

Intuitively, this means the synchrony of a group is dependent on the synchrony of the dyad most in-sync over the entire session. This is in contrast to selecting the maximum synchrony for each moment in time, dubbed “time maximum”, expressed by

$$Sync_{tmax}(P) = \frac{1}{s}\sum_{t=1}^{s} \max_{p,q\in P, p\neq q} m_{pt}\, m_{qt}$$

Note the difference in the orderings of the maximum operation and the sum. This second expression produces an average of the maximum value for each moment in time, allowing the dyad described as “in-sync” to vary potentially as often as every sample.

However, merely taking the maximum only takes one pair in a triad into account, meaning two dyads do not matter at all. Perhaps there is a middle ground between all dyads mattering equally and one dyad completely dominating any moment. Motivated by this intuition, we also investigate a family of functions S_α with a single parameter α that allows the function family to span between taking the average and the maximum of a set of numbers. The formal definition is

$$S_\alpha(x_1, \dots, x_n) = \frac{\sum_{i=1}^{n} x_i e^{\alpha x_i}}{\sum_{i=1}^{n} e^{\alpha x_i}}$$

This function takes an weighted average of the given values with weight given by the exponential of the value itself. A graph of $S_\alpha(0, x)$ for varying values of α is given in Figure 11.

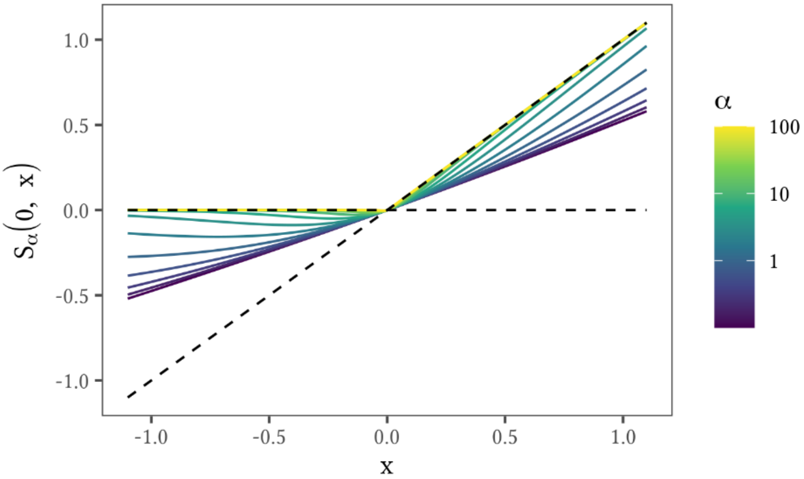

Figure 11. A graph of the smooth maximum aggregation functions and their values upon the arguments (0, x), represented by dotted lines. When alpha is small (near zero, dark purple), the function performs like an average, but when alpha is large (light green and yellow), it performs like a maximum.

Replacing maximum with a smooth maximum parameterized by some alpha provides two more measures of synchrony, $Sync_{sessm}$ for session-smooth-max and $Sync_{tsm}$ for time-smooth-max. In figure 12, we compare the measures defined by matching an aggregation level (time or session) and an aggregation function (max or smooth maximum), in addition to the $Sync_{avg}$ defined in section 3.5.3. We compare these measures on their ability to distinguish between synchrony and pseudosynchrony. The y-axis represents a logit scale of the percentile of real synchrony within the pseudosynchrony distribution. A high y-value indicates that real synchrony was larger than most pseudosynchronous values, while a low y-value indicates real synchrony was smaller than most pseudosynchronous values.

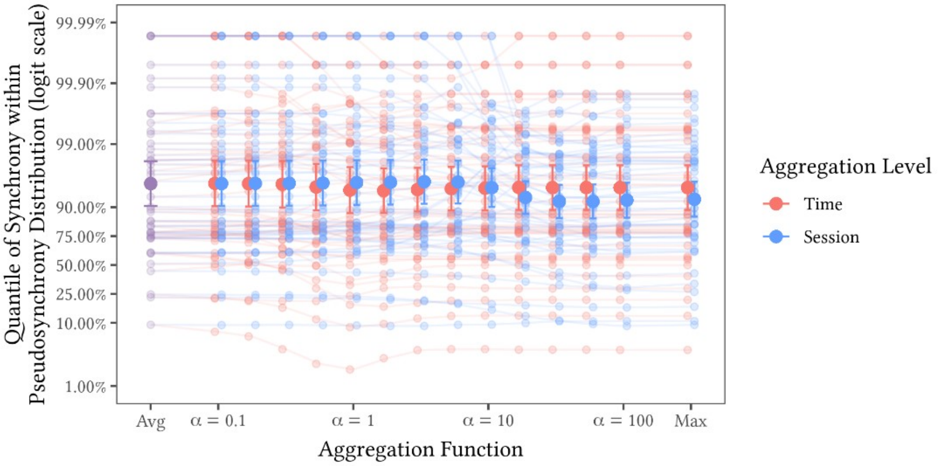

Figure 12. Mild differences in sensitivity to synchrony depending upon synchrony measure. Along the x-axis are the different aggregation functions, consisting of a spectrum with endpoints at average (left) and maximum (right) and smooth maximum in between. Large opaque dots indicate the mean, error bars indicate 95% confidence intervals of the mean. Groups of transparent dots connected by lines represent individual sessions.

Overall, the measures are similar, but there was a reduction in sensitivity on the maximum side of both measures. This drop happens at different values for the two different aggregation levels because the smooth maximum function is scale dependent. Note that merely comparing the error bars does not effectively capture differences, because the values across measures are very similar. In order to make some statements towards which measures are most effective, comparisons are made using paired t-tests. We do not perform multiple comparison correction due to the very exploratory nature of these tests. For this section we only report comparisons involving either $Sync_{avg}$, $Sync_{tmax}$, or $Sync_{sesmax}$, and we do not report comparisons between different parameters using the smoothmax function. These other comparisons, as well as exact p-values, are reported in the supplemental material.

Synchrony was better distinguished using $Sync_{avg}$ compared to $Sync_{sessm}$ with alpha values of 31, 56, and 100 (each $p < 0.05$), but performed worse compared to $Sync_{sessm}$ with alpha values of 0.1, 0.17, 0.31, 0.56, and 1 (each $p < 0.05$). $Sync_{tmax}$ was significantly worse than only $Sync_{\text{tsm}}$ with an alpha value of 1.7, ($p<0.05$). $Sync_{sesmax}$ distinguished synchrony better than $Sync_{sessm}$ with alpha values of 56 ($p < 0.01$) and 100 ($p < 0.001$) but worse than $Sync_{sessm}$ with alpha values between 0.1 and 10 (each $p < 0.05$).

In sum, the five different measures of synchrony produce different result, with methods more like maximum performing slightly worse. It is also worth noting that the variance between sessions dwarfs the variance between measures. However, these investigations are not just about practical considerations in creating sensitive synchrony measures, but also spurring new ways to test between theories of synchrony.

5 DISCUSSION

In section 3.6, we listed two hypotheses and five research questions. H1 was supported with statistically significant results, but H2 was not. We found evidence related to research questions RQ1, RQ4, and RQ5, but did not find results related to RQ2 or RQ3.

5.1 Synchrony

The immediate results of the study are further confirmation that synchrony occurs among participants in triads in virtual reality (H1), and that virtual location can have a significant effect on synchrony (RQ1). There was no evidence of differences in synchrony based upon task (RQ2), and no evidence that synchrony was related to performance (H2).

Our result for H1 follows from threads in previous work that virtual stimuli tend to elicit responses similar to real stimuli [47], and synchrony is a common and pervasive feature of human interaction [14]. We corroborate the results of both Wilkins and Nwogu [59] and Dale and colleagues [12] that synchrony exists among pairs within triads.

To explain on the causes of the difference in locations, we speculate two alternative explanations for the effect of location on synchrony. First, rather than the room signaling conformity due to the items within the space, it could be due to the space itself. High-ceiling buildings perhaps convey a feeling of smallness, encouraging a reduction of risk. Second, it is worth noting the conference room looked out over a downtown scene and triggered some participants' fear of heights. We are unsure how this increase in arousal may have influenced synchrony.

In regard to task, the lack of effect is similar to results reported in Sun and colleagues [54]. Their manipulation between competitive and cooperative conditions did not have an effect on synchrony. We interpret this based on the mathematical formulation of synchrony used in both studies, i.e., motion at nearly the same time. Even in highly competitive situations, such as an intense argument or a basketball game, interactants will exhibit synchronized motion.

The relationship of synchrony to performance in this study was not present with the sample size we ran. While a study with more statistical power may find a connection, we also would like to point to work by Ashenfelter and colleagues [3] regarding symmetry building and breaking in a conversation. In short, they note that the progress of a conversation consists symmetry building in coming to a common understanding and symmetry breaking as each participant takes time leading the conversation.

In addition to these results being findings on their own, they also demonstrate the future work necessary for a comprehensive theory of synchrony.

5.2 Exploratory Analysis of Synchrony

In the exploratory analyses we performed based on speaking data, we examined three research questions. First, we investigated the relationship of turn taking and synchrony. While we did not find any difference between transitioning and non-transitioning dyads (RQ3), we did find that the non-speaker dyad was more synchronous than speaker dyads (RQ4). We also explored the relationship of gaze to synchrony and found the surprising result that dyads were more synchrony when not looking at each other (RQ5).

Three possible explanations come to mind regarding the more synchronous listener dyads. Based on a conceptualization of synchrony as mutual interaction [21], one may claim that the two listening participants must have been interacting more with each other than with the speaker due to their higher synchrony. While it is true that listeners can and do communicate when not speaking, it seems unrealistic to claim that the listener is interacting more with the other listener than with the speaker, so we do not give much weight to this explanation. The second explanation conceptualizes synchrony to a type of shared attention [32]. If the listeners both see the speaker, and the speaker affects the listeners in roughly similar ways, then there will be a correlation in any kind of behavior between listeners. The third explanation is that this difference in synchrony is merely due to motion of the speaker. When the speaker talks, they may be more likely to move their head, and so be relatively high in terms of motion. However, because the listeners are not talking and so will have relatively low motion, they correlate negatively with the speaker and positively with each other. We judge the second and third explanations to have some weight and should be considering further.

We also found evidence indicating participants were more synchronous when they were not looking at each other (RQ5). This result seems to fly in the face of a theory of synchrony as mutual attention. However, there is a confounding factor that motion functions as both the value determining synchrony and a correlate of gaze behavior itself. This ambiguity may encourage explorations of other measures of synchrony.

Finally, we consider the exploration of triadic synchrony as a measure. While there is some evidence that $Sync_{sessm}$ with a small alpha value may outperform Sync_{avg}, we interpret this result with caution because these are p-values uncorrected for multiple comparisons. When also taking into consideration the simplicity of $Sync_{avg}$ as well as its presence in previous work, we still judge $Sync_{avg}$ to be the best measure of group synchrony for the time being. However, there is more to this exploration than just the question of the best measure.

5.3 Implications for Practice

Designers of virtual environments should continue to be aware that virtual environments are not merely backdrops to activity, they inform and constrain activities that take place inside them. It has been known that virtual people elicit realistic responses from participants in VR [6]; it is reasonable to extend this to virtual environments as well. From this study alone, it is unclear what features in these locations drive the difference in communication we found.

5.4 Limitations and Future Work

Some limitations of this work include sample size, participant pool, and technical issues. We note the small sample size of this study, as only 42 people were involved across the 14 teams used in the dataset. Future work can more accurately measure effect size by extending to dozens of teams. One potential confound was that some participants had prior experience in virtual reality, a fact volunteered to the authors during the experiment debriefing. While all participants spent some time at the beginning of the experiment acclimating to VR and all were able to complete the design tasks, it is likely prior experience influenced nonverbal behavior in some way. In addition, much of the data was unable to be analyzed, either through experimenter error or technical issues. Future work should be able to pre-register these hypotheses and analysis paths to produce confirmatory evidence of this effect.

While we do report a difference between the garage and conference room, only two locations were used. Using only this data, we cannot find evidence for what difference between these two spaces led to the difference in synchrony. Future work can explore finer differences between the rooms such as color, size, ceiling height, items present, and time of day, include changes as minor as changing the position of one object by a few inches.

The extension of synchrony to larger groups leads to further questions. What situational factors (team composition, task type) and individual differences (personality, communication styles) lead to the differences in synchrony among dyads in a triad?

Finally, interaction is not limited to synchrony. It is possible to also take into account the participant's behavior relative to the environment, e.g., what items are the participants standing beside, looking at, or interacting with. While this change would add complexity to the experiment design and analysis, it would help ground the computationally-heavy measurements with the semantics of human communication.

6 CONCLUSION

In this study, we have explored synchrony among triads in virtual reality. The results we have presented corroborate previous work that synchrony occurs in VR and among members of triad, indicating that synchrony is still robust even in these boundary conditions. This continues to solidify the value of synchrony as a subject of study. We also propose new multi-participant measures of synchrony for use and for study. Human interaction is often more complex than dyadic interaction; providing tools extending synchrony into this space supports future work understanding multiple participants in a team.

We have found some evidence suggesting that virtual environments can affect team synchrony and more broadly team interaction, though the roots of this effect are unknown. Echoing the results of Bandura's reciprocal determinism of person, environment, and behavior [4], and echoing environmental psychology more broadly, the virtual environment now becomes an opportunity for designers of VR content. We also have evidence contributing towards the understanding of what conditions synchrony effectively captures a meaningful nonverbal aspect of communication. The comparison between speaking and non-speaking dyads pits intuitive expectations of synchrony against the measures in common use, and the resolution of this expectation will help push the study of synchrony forward. Furthermore, the extension of the study of synchrony from dyads to triads may tease apart subtle differences in measurement and theory and set the stage for a larger insight on the nature of synchrony. Ultimately, this work better informs our understanding of human interaction, and affirms the long road ahead in future work.

ACKNOWLEDGMENTS

This work was supported by the Hasso Plattner Design Thinking Research Program grants during 2019-2020. The authors would like to acknowledge Will Gutzman, Dawit Gebrekidan, Jabreea Johnson, Siddhanth Chandra, Talia Weiss, and Tobin Asher for their assistance in this project.

REFERENCES

[1] Allsop, J. S., Vaitkus, T., Marie, D., & Miles, L. K. (2016). Coordination and collective performance: Cooperative goals boost interpersonal synchrony and task outcomes. Frontiers in Psychology, 7(SEP), 1–11. https://doi.org/10.3389/fpsyg.2016.01462

[2] Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. Journal of Personality and Social Psychology, 63(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

[3] Ashenfelter, K. T., Boker, S. M., Waddell, J. R., & Vitanov, N. (2009). Spatiotemporal symmetry and multifractal structure of head movements during dyadic conversation. Journal of Experimental Psychology: Human Perception and Performance, 35(4), 1072–1091. https://doi.org/10.1037/a0015017

[4] Bandura, A. (1978). The Self System in Reciprocal Determinism. American Psychologist, 33(4). Retrieved from https://www.uky.edu/~eushe2/Bandura/Bandura1978AP.pdf

[5] Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. Journal of Nonverbal Behavior, 12(2), 120–138. https://doi.org/10.1007/BF00986930

[6] Blascovich, J., & Bailenson, J. (2011). Infinite Reality.

[7] Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology Jim. Psychological Inquiry, 13(2), 103–124. https://doi.org/10.1366/0003702934066703

[8] Cheryan, S., Meltzoff, A. N., & Kim, S. (2011). Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. Computers and Education, 57(2), 1825–1835. https://doi.org/10.1016/j.compedu.2011.02.004

[9] Christy, K. R., & Fox, J. (2014). Leaderboards in a virtual classroom: A test of stereotype threat and social comparison explanations for women’s math performance. Computers and Education, 78, 66–77. https://doi.org/10.1016/j.compedu.2014.05.005

[10] Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. Developmental Science, 17(6), 1003–1011. https://doi.org/10.1111/desc.12193

[11] Condon, W. S., & Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior patterns. The Journal of Nervous and Mental Disease, Vol. 143, pp. 338–347. https://doi.org/10.1097/00005053-196610000-00005

[12] Dale, R., Bryant, G. A., Manson, J. H., & Gervais, M. M. (2020). Body synchrony in triadic interaction. Royal Society Open Science, 7(9). https://doi.org/10.1098/rsos.200095

[13] de Mendonça, J. S., Cossette, L., Strayer, F. F., & Gravel, F. (2011). Mother-Child and Father-Child Interactional Synchrony in Dyadic and Triadic Interactions. Sex Roles, 64(1), 132–142. https://doi.org/10.1007/s11199-010-9875-2

[14] Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal Synchrony : A Survey Of Evaluation Methods Across Disciplines.

[15] Dey, A., Piumsomboon, T., Lee, Y., & Billinghurst, M. (2017). Effects of sharing physiological states of players in collaborative virtual reality gameplay. Conference on Human Factors in Computing Systems - Proceedings, 2017-May, 4045–4056. https://doi.org/10.1145/3025453.3026028

[16] Dotsch, R., & Wigboldus, D. H. J. (2008). Virtual prejudice. Journal of Experimental Social Psychology, 44(4), 1194–1198. https://doi.org/10.1016/j.jesp.2008.03.003

[17] Driskell, J. E., Salas, E., & Driskell, T. (2018). Foundations of teamwork and collaboration. American Psychologist, 73(4), 334–348. https://doi.org/10.1037/amp0000241

[18] Eris, O., Bergner, D., Jung, M., & Leifer, L. (2006). ConExSIR: A dialogue-based framework of design team thinking and discovery. Studies in Computational Intelligence, 30, 329–344. https://doi.org/10.1007/978-3-540-34353-0_19

[19] Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. Vision Research, 51(17), 1920–1931. https://doi.org/10.1016/j.visres.2011.07.002

[20] Freeman, G., & Adkins, A. (2020). My Body , My Avatar : How People Perceive Their Avatars in Social Virtual Reality. 1–8.

[21] Galbusera, L., Finn, M. T. M., Tschacher, W., & Kyselo, M. (2019). Interpersonal synchrony feels good but impedes self-regulation of affect. Scientific Reports, 1–12. https://doi.org/10.1038/s41598-019-50960-0

[22] Gaziv, G., Noy, L., Liron, Y., & Alon, U. (2017). A reduced-dimensionality approach to uncovering dyadic modes of body motion in conversations. PLoS ONE, 12(1). https://doi.org/10.1371/journal.pone.0170786

[23] Goldschmidt, G. (2016). Linkographic Evidence for Concurrent Divergent and Convergent Thinking in Creative Design. Creativity Research Journal, 28(2), 115–122. https://doi.org/10.1080/10400419.2016.1162497

[24] Gordon, I., & Feldman, R. (2008). Synchrony in the Triad: AMicrolevel Process Model of Coparenting and Parent-Child Interactions. Family Process, 47(4), 465–479.

[25] Gordon, I., Gilboa, A., Cohen, S., & Kleinfeld, T. (2020). The relationship between physiological synchrony and motion energy synchrony during a joint group drumming task. Physiology and Behavior, 224(May), 113074. https://doi.org/10.1016/j.physbeh.2020.113074

[26] Grammer, K., Honda, M., Juette, A., & Schmitt, A. (1999). Fuzziness of nonverbal courtship communication unblurred by motion energy detection. Journal of Personality and Social Psychology, 77(3), 487–508. https://doi.org/10.1037/0022-3514.77.3.487

[27] Gumilar, I., Sareen, E., Bell, R., Stone, A., Hayati, A., Mao, J., … Gupta, A. (2020). A comparative study on inter-brain synchrony in real and virtual environments using hyperscanning. Computers & Graphics. https://doi.org/10.1016/j.cag.2020.10.003

[28] Hale, J., Ward, J. A., Buccheri, F., Oliver, D., & Hamilton, A. F. D. C. (2019). Are You on My Wavelength? Interpersonal Coordination in Dyadic Conversations. Journal of Nonverbal Behavior, (0123456789). https://doi.org/10.1007/s10919-019-00320-3

[29] Harrist, A. W., & Waugh, R. M. (2002). Dyadic synchrony: Its structure and function in children’s development. Developmental Review, 22(4), 555–592. https://doi.org/10.1016/S0273-2297(02)00500-2

[30] Herrera, F., Oh, S. Y., & Bailenson, J. N. (2020). Effect of behavioral realism on social interactions inside collaborative virtual environments. Presence: Teleoperators and Virtual Environments, 27(2), 163–182. https://doi.org/10.1162/PRES_a_00324

[31] Jablokow, K. W., Vora, A., Henderson, D. A., Bracken, J., Sonalkar, N., & Harris, S. (2019). Beyond Likert scales: Exploring designers' perceptions through visual reflection activities. ASEE Annual Conference and Exposition, Conference Proceedings.

[32] Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. Journal of Experimental Psychology: General, 146(4), 569–576. https://doi.org/10.1037/xge0000271

[33] kasenvr. (2020). /kasenvr/project-athena: Vircadia open source metaverse engine, based on the High Fidelity Virtual Reality Platform. Retrieved October 9, 2020, from Github Repository website: https://github.com/kasenvr/project-athena

[34] Latoschik, M. E., Kern, F., Stauffert, J. P., Bartl, A., Botsch, M., & Lugrin, J. L. (2019). Not Alone Here?! Scalability and User Experience of Embodied Ambient Crowds in Distributed Social Virtual Reality. IEEE Transactions on Visualization and Computer Graphics, 25(5), 2134–2144. https://doi.org/10.1109/TVCG.2019.2899250

[35] Lee, K. M. (2004). Presence, Explicated. Communication Theory, 14(1), 27–50. https://doi.org/10.1093/ct/14.1.27

[36] Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. Computer Mediated Communication, 3(2).

[37] Mccall, C., Bunyan, D. P., Bailenson, J. N., Blascovich, J., & Beall, A. C. (2009). Leveraging Collaborative Virtual Inter-Population Research on Persuation in a Classroom Setting. Presence: Teleoperators and Virtual Environments, 18(5), 361–369.

[38] McVeigh-Schultz, J., Kolesnichenko, A., & Isbister, K. (2019). Shaping pro-social interaction in VR an emerging design framework. Conference on Human Factors in Computing Systems - Proceedings, 1–12. https://doi.org/10.1145/3290605.3300794

[39] McVeigh-Schultz, J., Segura, E. M., Merrill, N., & Isbister, K. (2018). What's it mean to "be social" in VR?: Mapping the social VR design ecology. DIS 2018 - Companion Publication of the 2018 Designing Interactive Systems Conference, 289–294.

[40] Miller, M. R. (2021). dddr: Three-Dimensional Vector Operations in R. Retrieved March 15, 2021, from https://github.com/MrMallIronmaker/dddr

[41] OBS Studio Contributors. (2020). Open Broadcaster Software. Retrieved January 28, 2020, from https://obsproject.com/

[42] Oh Kruzic, C., Kruzic, D., Herrera, F., & Bailenson, J. (2020). Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. Scientific Reports, 10(1), 1–23. https://doi.org/10.1038/s41598-020-76672-4

[43] Paris, M. M., Carter, B. L., Traylor, A. C., Bordnick, P. S., Day, S. X., Armsworth, M. W., & Cinciripini, P. M. (2011). Cue reactivity in virtual reality: The role of context. Addictive Behaviors, 36(7), 696–699. https://doi.org/10.1016/j.addbeh.2011.01.029

[44] Purvis, C. K., Jones, M., Bailey, J. O., Bailenson, J., & Taylor, C. B. (2015). Developing a novel measure of body satisfaction using virtual reality. PLoS ONE, 10(10), 1–19. https://doi.org/10.1371/journal.pone.0140158

[45] Ramseyer, F., & Tschacher, W. (2008). Synchrony in Dyadic Psychoterapy Sessions. Simultaneity: Temporal Structures and Observer Perspectives., 329–347.

[46] Ramseyer, F., & Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: coordinated body movement reflects relationship quality and outcome. Journal of Consulting and Clinical Psychology, 79(3), 284–295. https://doi.org/10.1037/a0023419

[47] Reeves, B., & Nass, C. I. (1996). The Media Equation: How people treat computers, television, and new media like real people and places. Chicago, IL, US: Center for the Study of Language and Information.

[48] Roberts, D., Wolff, R., Otto, O., & Steed, A. (2003). Constructing a Gazebo: Supporting Teamwork in a Tightly Coupled, Distributed Task in Virtual Reality. Presence: Teleoperators and Virtual Environments, 12(6), 644–657. https://doi.org/10.1162/105474603322955932

[49] Rosedale, P. (2019). Updates, and a New Beginning. Retrieved from High Fidelity Blog website: https://www.highfidelity.com/blog/updates-and-a-new-beginning

[50] Salas, E., Reyes, D. L., & McDaniel, S. H. (2018). The science of teamwork: Progress, reflections, and the road ahead. American Psychologist, 73(4), 93–600. https://doi.org/10.1037/amp0000334

[51] Schroeder, R., Steed, A., Axelsson, A. S., Heldal, I., Abelin, Å., Wideström, J., … Slater, M. (2001). Collaborating in networked immersive spaces: As good as being there together? Computers and Graphics (Pergamon), 25(5), 781–788. https://doi.org/10.1016/S0097-8493(01)00120-0

[52] Slater, M., Spanlang, B., Sanchez-Vives, M. V., & Blanke, O. (2010). First person experience of body transfer in virtual reality. PLoS ONE, 5(5), 1–9. https://doi.org/10.1371/journal.pone.0010564

[53] Steed, A., Slater, M., & Sadagic, A. (1999). Leadership and Collaboration in Shared Virtual Environments University College London Adrian Bullock , Jolanda Tromp University of Nottingham. 3–6. https://doi.org/10.1109/VR.1999.756941

[54] Sun, Y., Shaikh, O., & Won, A. S. (2019). Nonverbal synchrony in virtual reality. 1–28. https://doi.org/10.1371/journal.pone.0221803

[55] Tarr, B., Slater, M., & Cohen, E. (2018). Synchrony and social connection in immersive Virtual Reality. Scientific Reports, 8(1), 1–8. https://doi.org/10.1038/s41598-018-21765-4

[56] Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., & Brooks, F. P. (1999). Walking > walking-in-place > flying, in virtual environments. Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '99, 359–364. https://doi.org/10.1145/311535.311589

[57] Walker, M. B., & Trimboli, C. (1984). The role of nonverbal signals in co-ordinating speaking turns. Journal of Language and Social Psychology, 3(4), 257–272. https://doi.org/10.1177/0261927X8400300402

[58] Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., … Yutani, H. (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686.

[59] Wilkins, N., & Nwogu, I. (2020). Analyzing the Extent of Rapport in Groups of Triads Via Interactional Synchrony. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020-Octob, 234–240. https://doi.org/10.1109/SMC42975.2020.9283474

[60] Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. Psychological Science, 20(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

[61] Won, A. S., Perone, B., Friend, M., & Bailenson, J. N. (2016). Identifying Anxiety Through Tracked Head Movements in a Virtual Classroom. Cyberpsychology, Behavior, and Social Networking, 19(6), 380–387. https://doi.org/10.1089/cyber.2015.0326

[62] Yaremych, H. E., & Persky, S. (2019). Tracing physical behavior in virtual reality: A narrative review of applications to social psychology. Journal of Experimental Social Psychology, 85 (April), 103845. https://doi.org/10.1016/j.jesp.2019.103845

[63] yozlet. (2020). yozlet/interface: Test platform for various render and interface tests for next-gen VR system. Retrieved October 9, 2020, from Github Repository website: https://github.com/yozlet/interface

Received October 2020, revised April 2021 and July 2021; accepted July 2021.